

182066

STIC-Biotech/ChemLib

---

From: Chernyshev, Olga  
Sent: Monday, March 13, 2006 4:38 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/736,936, sequence search request

Please search SEQ ID NO: 2 in regular and pending databases.  
Thank you very much!

*Olga N. Chernyshev, Ph.D.*  
*AU 1649*  
*REM 3C89*  
*2-0870*  
*mail 4C70*

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: 3-14-06  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 10:48:24 ; Search time 44 Seconds

(without alignments)  
1511.041 Million cell updates/sec

Title: US-10-736-936-2

Perfect score: 3598  
Sequence: 1 MQONQHNTAEKAPSENCK.....SLNKNGHPVPSAGADESHTIC 691

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2351	65.3	689	2 UC7286	liver-specific org
2	1411.5	39.2	670	2 UC7616	organic anion tran
3	1396.5	38.8	670	2 A49580	mediates transport
4	1334.5	37.1	670	2 UC7581	organic anion-tran
5	940.5	26.1	643	2 A41120	prostaglandin tran
6	894	24.8	682	2 UC7385	multispecific orga
7	723	20.1	809	2 T16448	hypothetical prote
8	686.5	19.1	1451	2 T16388	hypothetical prote
9	630	17.5	758	2 T26595	hypothetical prote
10	618.5	17.2	690	2 T27357	hypothetical prote
11	609	16.9	674	2 T21217	hypothetical prote
12	527	14.6	482	2 T17250	hypothetical prote
13	495.5	13.8	655	2 T16538	hypothetical prote
14	178	4.9	431	2 B81254	probable transmem
15	178	4.9	507	2 B84616	hypothetical prote
16	173.5	4.8	403	2 B84063	multidrug resistan
17	171.5	4.8	746	2 T05899	hypothetical prote
18	170.5	4.7	469	2 H97064	probable sugar-pro
19	160	4.4	487	2 B97884	probable tartrate
20	159	4.4	410	2 B84998	hypothetical prote
21	158	4.4	606	2 T27072	hypothetical prote
22	157.5	4.4	423	2 S74046	probable sugar tra
23	156	4.3	461	2 H64636	proline/betaine tr
24	153.5	4.3	425	2 AC2909	MFS permease limpo
25	149	4.1	613	2 SC7887	probable membrane
26	144	4.0	436	2 A81443	probable transmem
27	143.5	4.0	452	2 A71877	proline/betaine tr
28	143	4.0	471	2 H90502	phosphate transpor
29	142	3.9	457	2 D71144	hypothetical prote

30	138.5	3.8	541	2 T40518	major facilitator
31	138.5	3.8	567	2 A71619	membrane transport
32	137.5	3.8	422	2 H69839	multidrug resistan
33	137.5	3.8	613	2 T27077	hypothetical prote
34	137	3.8	433	2 B84566	hypothetical prote
35	137	3.8	513	2 T12997	hypothetical prote
36	135.5	3.8	474	2 B83719	multidrug resistan
37	134.5	3.7	457	2 H85059	probable sugar tra
38	134.5	3.7	464	2 F69587	L-arabinose transp
39	134.5	3.7	522	2 T41604	probable membrane
40	134	3.7	391	2 D71676	bicyclomycin resis
41	134	3.7	419	2 B90446	permease imported
42	133.5	3.7	489	2 D85434	hypothetical prote
43	133	3.7	387	2 D86640	multidrug efflux t
44	133	3.7	429	1 B41902	arsenical pump mem
45	133	3.7	497	2 S66834	probable membrane

ALIGNMENTS

RESULT 1  
UC7286  
liver-specific organic anion transporter-1 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 03-Jul-2004  
C/Accession: UC7286  
R:Ogura, K.; Choudhuri, S.; Klaassen, C.D.  
Biochem. Biophys. Res. Commun. 272, 563-570, 2000  
A>Title: Full-length cDNA cloning and genomic organization of the mouse liver-specific  
A/Reference number: UC7286  
A/Accession: J07286  
A/Molecule type: mRNA  
A/Residues: 1-689 <OGU>  
A/Cross-references: UNIPROT:Q9JUL3; UNIPARC:UPI0000023434; DDBJ:AB031959  
C/Genetics:  
A/Gene: lat-1  
C/Keywords: glycolysis; transmembrane protein

Query Match 65.3%; Score 2351; DB 2; Length 689;  
Best Local Similarity 63.8%; Pred. No. 4.8e-159;  
Matches 442; Conservative 97; Mismatches 144; Indels 10; Gaps 6;

QY	1	MQONQHNTAEKAPSENCKRYNGKLMFLALSLSFATKTLGATIMKSIHRRRE	60
DB	1	MQOTQHPSKA--AQPRLSEKTRHCDGFRIFLALSLFSYICKALGGVIMKSSITQIRRD	58
QY	61	ISSSLVGFIDGSPFIGNLIVIVFVSFYSGKLRPKLIGICRFIMGIGVLTALPHRFMGY	120
DB	59	IFSSISGLIDGSPFIGNLIVIVFVSFYSGKLRPKLIGICRFIMGIGVLTALPHRFMGY	118
QY	121	YYSKRTNIDSSNSTSTLCLINQILSLNRPSPFIVKGLKESGSMYIVFVGNL	180
DB	119	YVATENDISSLNST---LTCIVNQTSTLTGSPFIMKGEKGSNSTMYIVVLMGNL	175
QY	181	RKIGETPIVPLGLSYDDPFAKSGHSLVYGINATAMIGPIIGFTLGSFSMYVDIGVY	240
DB	176	RKIGETPIVPLGLSYDDPFAKSGHSLVYGINATAMIGPIIGFTLGSFSMYVDIGVY	235
QY	241	DLSIRVITPDARWGWAMWLGFTVNGLLCTISIPFFLPKLPKKSQKRKNSASLHVL	295
DB	236	DLSIRVITPDARWGWAMWLGFTVNGLLCTISIPFFLPKLPKKSQKRKNSASLHVL	295
QY	301	TNDEKQDTANLNQGNITKAVTGFQSKSLITNPLVYMPFULTLLQVSSYIGAFITYF	360
DB	296	TNDEKQDTANLNQGNITKAVTGFQSKSLITNPLVYMPFULTLLQVSSYIGAFITYF	355
QY	361	KVVEOYQGPSSKANILGVITIPFASGPFAGYIKKFKNTVGIASFATAMSLS	420
DB	356	KVVEOYQGPSSKANILGVITIPFASGPFAGYIKKFKNTVGIASFATAMSLS	415
QY	421	FTLLVFFLICENKSVAGLTMTYDGNPVTSHRDVPLSYCNSDCNDESGMEPVCGANGIT	480
DB	421	FTLLVFFLICENKSVAGLTMTYDGNPVTSHRDVPLSYCNSDCNDESGMEPVCGANGIT	480

Db 410 FYLSYFLLICENKAPAGLLTLTYDGMNPVDSHIDVPLVCNSPDCIDKNQWEPVGENGVT 475

Qy 481 YISPCLAGCKSSSGNKK--PIYFNCSCLEAVTGLONRYSAHLEGCPDDACTKTFPEV 533

Db 476 YISPCLAGCKSPFGDKGMNIRFYPDSCVSGSGFGKGNHSARLGECP--DCKTKTYFYI 534

Qy 539 AIDVLI,PSFALGCTSHMLVIKIVQPELKSIALAFGHSMLVRALGGLIAPYFGALIDTY 598

Db 535 TPQVYIISFPLALGISTIMLILIRSVQPELKSIALAFGHSMLVKTGLIAPYFGALIDRT 599

Qy 599 CIRKSTNNCGTRGSCRTYNSTSPFSRYVYLGSMLPVSSLVYIILYIMAKKYYQEKDINA 655

Db 595 CMKMSVTCGARGACGLYNSRLFGMIYVGLSLATKTPILILYVALIYMKRMRKRNDKI 654

Qy 659 SENG-SYMDENALIESLNKHK--FVPSAGADSET 689

Db 655 LENGKRFTEGNEPEPVNNNGYSCVPEDEKNSST 687

## RESULT 2

organic anion transport polypeptide 2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Dates: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004  
C:Accession: J07616  
R:Ogura, K.; Choudhuri, S.; Klaassen, C.D.  
Biochem. Biophys. Res. Commun. 281, 431-439, 2001  
A:Title: Genomic organization and tissue-specific expression of splice variants of mouse  
A:Reference number: J07616; MUID:21092843; PMID:11181066  
A:Accession: J07616  
A:Molecule type: mRNA  
A:Residues: 1-670 <OGU>  
A:Cross-references: UNIPROT:Q9EP96; UNIPARC:UPI00000271ER; DBJ:AB011814  
C:Comment: This protein with twelve transmembrane domains, glycosylation sites and proteo-  
C:Genetics:  
A:Gene: catp2  
C:Keywords: transmembrane protein  
F:21-39/Domain: transmembrane #status predicted <TM1>  
F:59-80/Domain: transmembrane #status predicted <TM2>  
F:87-105/Domain: transmembrane #status predicted <TM3>  
F:157-179/Domain: transmembrane #status predicted <TM4>  
F:192-214/Domain: transmembrane #status predicted <TM5>  
F:244-266/Domain: transmembrane #status predicted <TM6>  
F:317-339/Domain: transmembrane #status predicted <TM7>  
F:354-376/Domain: transmembrane #status predicted <TM8>  
F:387-409/Domain: transmembrane #status predicted <TM9>  
F:514-536/Domain: transmembrane #status predicted <TM10>  
F:550-571/Domain: transmembrane #status predicted <TM11>  
F:600-622/Domain: transmembrane #status predicted <TM12>

Query Match	39.2%	Score 1411.5	DB 2	Length 670
Best Local Similarity	43.2%	Pred. No. 2.2e-92		
Matches 290	Conservative 134	Mismatches 214	Indels 33	Gaps 12

Qy	27	IKHFLALSLSPKAKTGLIMKMSLIHERRFEELSSLVFIDSPFGLMLIVFVSY	86
Db	19	IKHFLALTCAYVSKSLSTGYMSMLTQHERPGIPTSVMGLINSFEGNLLIIFVSY	78
Qy	87	FGSKLHRPKLIGIGCFIMGIVLTAIPHFMGYRYRSKEFNIDSSSENSTSLTCLINQ	146
Db	79	FGYKLRHPIMIGVGCAMVGLGCFLLISIPHLTGRAYEY--ETTLIPTSMLSNPFCTENR	133
Qy	147	ILSLNRPSPRIVEKGLCKESSGYMMIYVGMGLMLGIGETPIVPLGSLYIDDPKKEGHS	206
Db	137	TQTL---KLPQDTECVKKNKSLMMIYVLVGNILKMGSTPIIMPGLSISYIDPAKSENRP	199
Qy	207	LYGLIINALAMIGPIILIGFTGLSLFSKMYVDIGYDLSTRITPTDSRVGVGMMLPLVSG	266
Db	194	LYIGILLETGMTIGPLGLGLSSCANIYVDTSVMTDLITTPTRVWGVAMMIGPLVCA	255
Qy	267	LBSIISISPPPLPQTPNPKQÖKERKASLSLHVLFTNDEKQDTAULTNÖKNITKAVTGFF	322
Db	254	GVNVLISIPFFPPKTLRKGLQNDGDG---ENAKERKREKIKENNRDITD---FF	306

[illegible]

### RESULT 3

mediates transport of organic anion - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: A49580  
R:Jacquemin, E.; Hagenbuch, B.; Stieger, B.; Wolkoﬀ, A.W.; Meier, P.J.  
Proc. Natl. Acad. Sci. U.S.A. 91, 133-137, 1994  
A:Title: Expression cloning of a rat liver Na(+)-independent organic anion transporter.  
A:Reference number: A49580; MUID:94105118; PMID:8278355  
A:Accession: A49580  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-670 <RSS>  
A:Cross-references: UNIPROT:P46720; UNIPARC:UPI0000135446; GB:IL19031; NID:G410310; PIDD  
C:Genetics  
A:Gene: catp

Query Match	38.8%;	Score 1396.5;	Length 670;
Best Local Similarity	42.9%;	Pred. No. 2.5e-91;	
Matches 289;	Conservative 125;	Mismatches 231;	Indels 29;
			Gaps 11

```

QY      2  RYNGCKMLPAAALSSLFIAKTGATIMSSIIHIERREISSIVGFDGFEIGNLVI  81
Db      14  RLRSKKNVFLSLSTCACTKLSLGYVMSMLQIERQDISTSVAGLINSFEIGNLFI  73
QY      82  VVVSRYGSKLHAPKPLIGICGFTMGIGCVLTALPHEFMGYTRYSKETNIDSENSTSLST  144
Db      74  VFVSRYFOTKLRHVRVVGIGICVMIGLGCCLMSLPHFPMRYEY--ETTISPGNLSNSFL  139
QY      142  CLINQILSNRAPREIVTGKGLCKESGYVMIYVFGNNMLARGIGETPIVPLGLSYIDDPAK  201
Db      132  CMENRRQTL---KPTDDPABCVKREKMSLMNLCVMNGNIRIGIEPTPIVPLGISYIEDPAK  188
QY      202  EGHSSLYGLILNAILAMIGPIIGFTGSLFSKATYDIGYVDSLIRIPTDSRWNGAWMLN  261
Db      189  SENSPLVTIGLMEKGKACGPIFGILLGSYCAQIYVDIGSVNDDLTLPSPDTRRWGAAMIG  248
QY      262  FIVSGLESIISLIPPEFLPOT--PNKPKQERKASLSLHVLFTNDEKDOT--ANLTNYQCKNI  318
Db      249  FLVCAGNILNITSLPEFFLPKALPKKQGD-----NVAYTKDKQVKEYGQGAABENULGI  304
QY      319  TKQNVGFPOSEKSLNPLNPLVFLTLTLOVSSYIGAFYVFKYVBOQYQPSKANIIL  378

```

```

Db      302 TND---FLTFMKRLPCNPITMLFILTSLVQVNGFIMKFTFLPKRYLEQOYGKSTAEKIFILI 358
      379 GTITIPFASGMLGGYIIKKFKLNTVGIKAKSCPTAVMSLSFYLLYFPIICENKSVAGL 438
      359 GYVSLPEPICGLYGLGFMKKFKITVKKAAVLAFLCLSVPEYLLFLCHFLMLTCDMAVAGL 418
Qy      439 TMTYDGNMVTSHRDVPLSYCNSDCDSQWEPVCGNNGITYIISPLACGKSSGKKKPY 498
      419 TTSYGVQHQHVLVESKVLADCNTRCSCSTNTWDPVCGDGVAYMSCLAGCKKFPVGTGN 478
Qy      499 IYFVNCSCLEVTGLQNRNYSALHGECPRDACCTRKYPFVAIOVNLFPSSALGSGTSHWL 558
      479 WYFQDCSCIGSG---NSSAVLGCLCKKPECANRLOVFLIITIIISFYSLTAIPGYNV 534
Qy      559 IYKIVQPELKSIALGFHSMVIRALGGILAPIYFGALIDTTCIKWSTNCGTRGSCRTYNS 618
      535 PLRCVSEKESKLGVLGHTFCIRVPAGIPAPVYFGALIDRTCLHMGTCLKGCGAGCMYDI 594
Qy      619 TSPSRVYLGSLSMRLRSS-LVLYIITLYAMKKKYQ---EKDINASENGSMDEANLESIN 674
      595 NSFRHLYLGLPIALRGSSYLPAPFILI--LMRKFPQPDIDSSATDHEMM-LGEEKSEH 651
Qy      675 KKHGFVPSAGADSE 688
      652 TDVHGSPOVENDGE 665

```

## RESULT 4

```

JCT581
organism anion-transporting polypeptide-5 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C:Accession: JCT581
R:Choudhuri, S.; Ogura, K.; Klaassen, C.D.
Biochem. Biophys. Res. Commun. 280, 92-98, 2001
A:Title: Cloning, expression, and ontogeny of mouse organic anion-transporting polypepti
A:Reference number: JCT581; PMID: 21092571; PMID:11162463
A:Contents: Kidney
A:Accession: JCT581
A:Molecule type: mRNA
A:Residues: 1-670 <CHO>
A:Cross-references: UNIPROT:Q99494; UNIPARC:UP100000279D2; GB:AF230701; GB:AF213260
C:Comment: This protein, an ortholog of rat, is a kidney-specific organic anion transport
C:Genetics:
A:Gene: oacp-5
C:Keywords: kidney; transmembrane protein; transport protein

```

```

Query Match      37.1%; Score 1334.5; DB 2; Length 670;
Best Local Similarity 42.2%; Pred. No. 6,3e-87;
Matches 278; Conservative 115; Mismatches 238; Indels 27; Gaps 10;

Qy      27 LKMFPLAALSPIAKTLGAIIMKSSIHIERRFEISSSLVGFIDSGFEIGNLLVIVFVY 86
      19 IKVFLIALIMAYISKILSGVMSTMLTQERQFNISTISVGLINSFEKGNLLVIVFVY 78
Qy      87 FGSKLHRLPLIGICGFMIGIGVLTALPHRFMGYRYRSKNTNIDSSNSTSTLSTCLNQ 146
      79 FGTKLHRLPMIGVGVAGVGLGCFIISLPHFLMGREY--ETTISPTSMNSNSPFLCVERK 136
Qy      147 ILSLNRASPEIYVKGKCLKESSGYMMIYVFMGMNLGIGETPIVPLAGLSIYIDPFAKGS 206
      137 SGTL---KPTODPACVCEIKSLMIYLVGNIRIGIGETPIIMPAGISITIEPFAKSNP 193
Qy      207 LYLGLILNLIAMIGPIIGTFLGSLFSKMYVDIGYVDSLITRITPTDSRWGVAGWMLNVLVSG 266
      194 LYIGILEVGMKIPILGILYLMGPFCANITYVDTSVNTDILTPTDTRWGMWIGFLVCA 253
Qy      267 LFSITISSIPFPLPOTPKPKQEKERKASLSLHVLFTNDEKDOQANLTLNQGKNTTKAVTGF 326
      254 GNVVLTSTSPFFFPPTLPBEGLDQNDGDEAKNK--EKGRDKAKENQD--TIKK---FF 306
Qy      327 QSFKSLITNPLVVMFVLLTLLOVSSYIGAFYVFKYVEQVQGPQSKANILLGITIPIF 386

```

```

Db      307 LMKMLPCNPITMLCVLTSVLQVNGVANIYKPKLBNHFGISTAKAVFLGLVYTPSV 366
Qy      387 ASGMLGGYIIKKFK--LNTVGIKAKSCPTAVMSLSFYLLYFPIICENKSVAGLMTYDG 444
      367 SAGVILISGIMKMLKITLKKAINALCLFMSCLLS--LCNFMLTCDTTPPIAGLTTSYSG 424
Qy      445 NNPTVSHRDVPLSYCNSDCDSQWEPVCGNNGITYIISPLACGKSSGKKKPYIYVNC 504
      425 IQOSFDMERKFLSDCNTRCNCLTKTWDPVCGNNGIAYMSPCLAGCEKSYGTGANMWFQNC 484
Qy      505 SCLBYTGLQNRNYSALHGECPRDACCTRKYPFVAIOVNLFPSSALGSGTSHWLYIKYIO 564
      485 SCIRSSG---NSSAVLGCLCKKPECANRLOVFLIITVFCCEFFSLATTIPGMVFLRCKK 540
Qy      565 PELKSIALGFHSMVIRALGGILAPIYFGALIDTTCIKWSTNCGTRGSCRTYNSPSRV 624
      541 SEKSLGIGLQAFMRFLFAGIPAPVYFGALIDRTCLHMGTCLKGEGGACRTYEVESFRL 600
Qy      625 YIGLSMLRVSLVLYIITLYAMKKKYQ-----EKDINASENGSMDEANLESIN 676
      601 YIGLPAALR-GSIIIPSEFFILRLKLOIPGDDISSEIHLAETKTEKSECTDMKS 657

```

## RESULT 5

```

A41120
prostaglandin transporter - rat
N:Alternate names: matrix F/G
C:Species: Rattus norvegicus (Norway rat)
C>Date: 27-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
C:Accession: A41120
R:Hakes, D.J.; Berzney, R.
Proc. Natl. Acad. Sci. U.S.A. 88, 6186-6190, 1991
A:Title: Molecular cloning of matrix F/G: a DNA binding protein of the nuclear matrix ch
A:Reference number: A41120; PMID:91296785; PMID:2068100
A:Accession: A41120
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-643 <HAK>
A:Cross-references: UNIPROT:000910; UNIPARC:UP10000135445; GB:M64862
A:Note: the authors did not translate the codons for residues 1-99 in this reference
R:kanai, N.; Lu, R.; Satriano, J.A.; Bao, Y.; Wolkoff, A.W.; Schuster, V.L.
Science 268, 866-869, 1995
A:Title: Identification and characterization of a prostaglandin transporter.
A:Reference number: A38955; PMID:95273959; PMID:7754369
A:Contents: annotation
C:Keywords: transmembrane protein

```

```

Query Match      26.1%; Score 940.5; DB 2; Length 643;
Best Local Similarity 32.3%; Pred. No. 5,4e-59;
Matches 215; Conservative 131; Mismatches 240; Indels 79; Gaps 15;

Qy      25 NGLKMFPLAALSPIAKTLGAIIMKSSIHIERRFEISSSLVGFIDSGFEIGNLLVIVFV 84
      29 SNIKVFLVCHGLLQCLLYSAVFXSLTTIEKRGLSSSSGLISLNRISNATLIPI 88
Qy      85 SYFGSKLHRLPLIGICGFMIGIGVLTALPHRFMGYRYRSKNTNIDSSNSTSTLSTCL 144
      89 SYFGSRVVRPMIIGIGLLAAGAVLTLPPLSPPYQYTSY---DGNSSYQTDI- 142
Qy      145 NQILSLNRASPEIYVKGKCLKESSG-----SYMMIYVFMGMNLGIGET 186
      143 -----CQKHFGALPPEKCHSTVEDTHKENSILMGWVVAQLLAGIGIV 185
Qy      187 PIVPLGSLYIDDPAKGHSYLYGLILNLIAMIGPIIGTFLGSLFSKMYVDIGYVDSLSTR 246
      186 PIQPGISYVDDPAPTSPLYISILFAIVGPAFGYLGSLVMLRIFVYGRVDTATYV 245
Qy      247 ITPDTSRWGVAGWMLNVLVSGLFSIISIPFPLPOTPKPKQEKERKASLSLHVLFTNDEK 306
      246 LSPGDPRIAGMWLGLLISGFLVTSLPFFPPPA---MSRAERSV---TAETRM 296
Qy      307 QTANLTNQGKNTTKAVTGFQSFKSLITNPLVVMFVLLTLLOVSSYIGAFYVFKYVEQ 366

```

Db 297 QTBEDKSRG--SLMDPIKRFPRILFLRLNPNLPLMLVVLVQCPTSSVYAGLSIFLAKFEKQ 355  
QY 367 YGQPSKANIILLGVITIPDIPASGMFLGVIIKKFPLANTVGIKAKESCPTAVNLSLFYLLYF 426  
Db 356 YGATAAANFLIGAVNLPLAALGLMFGGILMKRFVPLQTIIPRVATITITSMICVPLF 415  
QY 427 FLELENSVALMTYTDGNPNPVTSHRDVPLSYCSDNCDCDSQMPVCYGNNGITFYISPL 486  
Db 416 FMGCTSGAAVAV---YPPSTSSSIHPQOPPA--CRDDCSGPPSPFHPVGDNGEVVSCH 471  
QY 487 AGCR-----SSSGNKKPIVFYNVSCLEVTGLQNRNYSAHIGECPRDDACTRKFYFVAIOV 542  
Db 472 AGCSTNTVSSASKRPI--YLNCSC--VSG-----GSAQDRMLPHTLRALLPSI 518  
QY 543 LNFPSAL--GGTSH---VMLVYKIYQPELKLIALGFHSMVIRALGILAPFYFALITTT 598  
Db 519 PLISFPAALITACISHNPLVMVLRVNVODEKSFALGVQFLMKRLMLAPASISYGLLIDSS 578  
QY 599 CIKSTNNCTGRGSCRTYNSTSPFSRVYGLKLSMLRVSSLVVLVYIILYAKKKKYOEKDINA 658  
Db 579 CVRNNVLCSGRGACAYYDNDALNRRIYGLQMVYVYALGTLILLFISMWK--NREYSL 635  
QY 659 SENG 663  
Db 636 QENT 640

```

RESULT 6
JC7385
multispecific organic anion transporter, moatl - rat
C|Species: Rattus norvegicus (Norway rat)
C|Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C|Accession: JC7385
R|Nishio, T.; Adachi, H.; Nakagomi, R.; Tokui, T.; Sato, E.; Tanemoto, M.; Fujiwara, K.;
.: Innama, K.; Nunoki, K.; Matsuno, S.; Abe, T.
Biochem. Biophys. Res. Commun. 275, 831-838, 2000
A|Title: Molecular identification of a rat novel organic anion transporter moatl, which
A|Reference number: JC7385
A|Contents: Brain
A|Accession: JC7385
A|Molecule type: mRNA
A|Residues: 1-682 <NLS>
A|Cross-references: UNIPROT:O9JH13; UNIPARC:UPI000013544B; GB:AF169409
C|Comment: This protein, a transmembrane glycoprotein, transports prostaglandin D2, leu-
rostaglandin E1, E2, thromboxane B2, and ilprost.
C|Genetics:
A|Gene: moatl
A|Keywords: brain; glycolysis; transmembrane protein

```

[illegible]

```

Oy 295 SLHVLTEBKD---OTANULTQGNKITKNTYNGFFQSKSI----LTPNLYMVLVLTLL 347
Db 313 GSKGBEILSSCHEPLKXKQAGLPQIADPLT--VVQFIKVPYRLARTLRHPILFLLVLSQVC 370
Oy 348 QVSSYIGAFYVYKVKYEQOYQGPSSKANILLGVITPIIFASGMFAGYIILKFKPLNTVGI 407
Db 371 TSSWAGTATFLPEKFIEROPSTI5AFNILLGCLTPIPLAIGYIVGVVLVGRLLHLSMQC 433
Oy 408 AKESCFTAWMSLSFYLLLYEPILCENK5VAGLTMVTDGNNPVTSHRDVPLSYCNSDCNDE 467
Db 431 SALCLIGSILLCILLSLPLFFIGCSHTHIAGITQDL-GAQPGPS---LFPSCSPSCQS 485
Oy 468 SQMEPVQGNNG-ITYSPLCLAGC---SSGNKKPIYFNVCSCLEATYGLONRYSAHLG 522
Db 486 DDRNPVCDTSAYVEYTPPCBAGCTGRVQVQBALDKSQVEPTVCSC--VAG---NGTISAG 539
Oy 523 ECPRDACTCRKFYFPAIVLNLFP5ALGTSHWLKYIKVOPELKSALAGFSHMVIRAL 582
Db 540 SC--BSACSRVLVLPFLILSLGAAVASITHTPSPFLILRGKXEDKTLAQMCPMLRLV 597
Oy 583 GGIILAIYFQALIDTTCIKMSTNNCGTRGSCRTYNSTFSRBYVLGLSMLRVSLVLYIT 642
Db 598 AMMPSPVINGSAIDTTCVHMAL-TCGRABVACRYYDHLLRMRFTGLGFFKSGSLVCFAL 656
Oy 643 LIYAMKKKIQOEKDINNA5ENG5YMD 667
Db 657 VIALIRQSGREASTKATYKVSQDLQE 681

```

```

RESULT 7
T16448
hypochlorite protein F53B1.8 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C|Accession: T16448
R|Miller, N.
submitted to the EMBL Data Library, November 1995
A|Description: The sequence of C. elegans cosmid F53B1.
A|Reference number: Z18514
A|Accession: T16448
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-809 <M.L.>
A|Cross-references: UNIPROT:Q20702; UNIPARC:UP1000007995A; EMBL:U40953; NID:g1072248; FR
A|Experimental source: strain Bristol N2, clone F53B1
C|Genetics:
A|Gene: CESP.F53B1.8
A|Map position: X
A|Introns: 39/2; 66/3; 116/3; 214/1; 321/2; 405/1; 580/2; 643/3; 750/3

```

Query Match	20.1%	Score 723	DB 2	Length 809
Best Local Similarity	27.6%	Pred. No. 1,96-43		
Matches	202	Conservative 135	Mismatches 272	Indels 122
				Gaps 26
Qy	10	TAAQPSSEK----	KTRYCNGLMPLAALSLSEIANTLGAII----	MKSIIHIERREISS 63
Db	81	TESSQCGICKMRKMLQCGKQLPMIVL-LCYGSIQGLIVGLVPSAIIISIERRFEST 139		
Qy	64	SLVGFIDGSEFIGNLLIVFVSYPGSKLRPLIGCFINGIGCVLTALPHFMGYRR 1233		
Db	140	SHHGRIVQFDEFQYVLICIPVSYPGGRHSKPVVLVLAAGIACALGSLIFSAPHMSDSE-- 197		
Qy	124	SKETNIDSSNSNSTLCTLIHQIILSRAPSEIVGK-GCLKR-----	SGSYMIIVF-M 176	
Db	198	-----SASQSMGSGTCSIEHLQNDMTSEALQIQOSSNQDGHSSNNLYPIYLCF 256G		
Qy	177	GNNLRIGIGETPIPLGLASYIDDPAKGGHSLYLGIINLAMIIGPIIGFTLGSLSKMYVD 236		
Db	251	AHLRHIGIATPLETIGSVYIIDENVTAASSLFVGFYFAVFGALIGFLASVSLQVHTD 310		
Qy	237	IGYV----DSTRITPTDSRWYGAWMLNLFVSGLFSIISLPFFLQTPMKPKQEKAS 2933		

```

Db      311 FMHLPEQILKVSSESTDPWVGAMWLSPIAASFVGAVVLPASLP-----KVLPS 363
Qy      294 LSLHLEATNDE-----KQDTANLT-----NQGK----- 317
Db      364 LKMRHRLHDAAVANRRRTPECCMGPMGKNTALNSDAPADNFGSKAIKKEDYTTFFM 423
Qy      318 -----ITKNVGFQSGFSK-----ILNPLVYMFVLLTLQ 348
Db      424 LNKIFEITQASLLYSIPARGRGLMYKIWDVRHIPLAIYRILTNMLFM--IITSAMA 481
Qy      349 VSSYI--GAFYVFKVVEQYQGPSKANIILGVITPIFASGMPLGYIILKKEKNTVG 406
Db      482 ISLVVTTGASSFMSKLEHGFVSVPKSNALLIGCVAVPMAGMCMWTGGLVNHFRANSSK 541
Qy      407 IAKFSCFVAVMSLSFYLLYFFILCENKSVAGLTMTY--DGNP--VTSHRDVL-SYC 459
Db      542 MKLFAIGLIFLSLFFSPMY-LIYCPHAPLVGVDSAYPEFGSIPDYNATYSKDPBLMNSC 600
Qy      460 NSDQNCDESOEWBVC-----NNGITISPLACGKSS--GSKKPIVFNVCSCLEVTGLN 514
Db      601 NKOCTDPBSYRVCALDGDGROFTYSPCYAGCSTYSASLKE--YTSCSVP-TNTN 657
Qy      515 RNVSAAHAGCEPRDACTRKFFVFAIQLVNLFFSALGTSHWALIVKIYQBELKSLALGF 574
Db      658 RPRIVKKGVC--BPGCTELLMFLFLFAPLSFCFFAL-AVPIISVILRTVDYMERFALGI 714
Qy      575 HSNVIRALGGLIAPYFAGALIDTTCIKMSTNNGC-TRGSCRTYNSFSRVLGLSMLR 633
Db      715 QWIMIVIGTIPAPVLFGMFVDSICKYQSDSCSGSGCLMYSNQYLANLFLFTISGQ 774
Qy      634 VSLVLYITLI 644
Db      775 ITTLALIVL 785

```

## RESULT 8

T16388

hypothetical protein f47B1.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C/Accession: T16388

R/Pavello, T.

submitted to the EMBL Data Library, November 1995

A/Description: The sequence of C. elegans cosmid F47B1.

A/Reference number: Z16505

A/Accession: T16388

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1451 &lt;FAV&gt;

A/Cross-references: UNIPARC:UPI000017B9DF; EMBL:U09993; NID:G1055098; PID:G1055100; PIDN

C/Genetics:

A/Gene: CESP:F47B1.2

A/Introns: 50/1; 70/2; 108/3; 154/3; 249/3; 331/1; 522/2; 551/2; 627/2; 725/3; 773/1; 79

Query Match 19.1%; Score 686.5; DB 2; Length 1451;

Best Local Similarity 24.5%; Pred. No. 1.5e-40;

Matches 186; Conservative 134; Mismatches 267; Indels 177; Gaps 20;

```

Qy      29 MFLAALSLPIAKTIGAIIMKSSIIHIERRFEISSLVGTFIDSGFEIGNLIVFVSYFG 88
Db      6 FFLGASFVFPFLBAGISYINSVQNIHQFQWSSRTSGMIBASDFGYITPCVVFIAHFG 65
Qy      89 SKLHRPKLIGIGCFIMGIGVLTALPHF----- 117
Db      66 SKGNRRARWIGAGAILVISITLMASBNFLFPGGHDINTTEVSAKKPTAGOLASNSTLK 125
Qy      118 -----MGYYYSKETNID-----SENSTSTLSTCLIN 145
Db      126 QLLSYGLIDRMHMDYMLNLNEDDTPIEYQPIGRPKKGGGSGNSTYTIIDQMLN 185
Qy      146 QTL----- 155
Db      186 EALVALENTIANNTGNTLRSSLSLFIHRRTNTSTYDIOKIRISAAPAFPGCKLTNSLR 245

```

```

Qy      156 EIVKGCCKR--SGSYMTIVFMGNLRIGIEPRVPLGLSYTDDPAKGHSLYTGILN 213
Db      246 AVIKOSKCEQSTNSVPFLVFPFSLLLIGIRVPMSLVPLDDNIKKKSLPAYGALS 305
Qy      214 AINMIGPIIGFTGLSFSKMYVDIGYVDSLITRTPTDSRWVGMNLVLSGLFSLSS 273
Db      306 SIRVLGPIGMYAGSCNKFYTLNPN-----GLTPADPTWICAMMGFLGSLALFSS 361
Qy      274 IPEFLLPQTPNKRQKERRKASLSLHLETND-EKQDTANLTNQGKITKNVGFQSGFSK 332
Db      362 TMLFFPQ-----GKEGDS-----AVQLRDVHKKKKGVDEDRITMMLKDFAKSCKKV 412
Qy      333 LTNPLVYMFVLLTLQYSSYIGAFYVFKVVEQYQGPSKANIILGVITPIFASGMPL 392
Db      413 LSKRYMGSVLAGVCVLAFFKGIYVLPYLEHFGIPQYLVRHYAMFGVGFGLGVAT 472
Qy      393 GGYIILKFKLNTVGIKAFSCFTVAMSLFYLLYFFILCENKSVAGLTMTYDGNPVTSHR 452
Db      473 GGYVTKKLINGRAAMFVMLNSTLANCLYSGKIFIGCS-----IVNSIGNNRQNTYN 527
Qy      453 DVPLSYNSDQNCDESOEWBVCNNGITYISPLACGKSS--GSKKPIVFNVCSCLEVT 510
Db      528 FT--RECNSQSCENARLYPVCQDTGFAYFSPCHAGCREAMQYGSPPVLDFTSCQCAP-G 584
Qy      511 GLONRYSAAHAGCEPRDDAC-TRKFFVFAIQLVNLFFSALGTSHWALIVKIYQBELKS 569
Db      585 GVSKKPF-----C-ENKCTKSSVIFLTV-LPGSFVAGLVPMALMLKRSVPETSS 635
Qy      570 LALGFSNVRALGGLIAPYFAGALIDTTCIKMSTNNGC-TRGSCRTYNSFSRVLGLS 629
Db      636 LSLGLQMAVSLPGLTIPSLIMGLVIDAOLVMDKACNBARSGCSLYH-----P 684
Qy      630 SMLRVSLLVLYTL-----IYANKKKYQEKDINASNGSYMDS 667
Db      685 DRLKVMHLLVYVIRFVALITDYYWKH-----AKQLNIMDR 721

```

## RESULT 9

T26595

hypothetical protein y32F6B.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T26595

R/Barlow, K.

submitted to the EMBL Data Library, January 1998

A/Reference number: Z20242

A/Accession: T26595

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-756 &lt;WIL&gt;

A/Cross-references: UNIPROT:O62421; UNIPARC:UPI000007B1F3; EMBL:AL021475; PIDN:CAA16312.

C/Genetics:

A/Gene: CESP:Y32F6B.1

A/Map position: 5

A/Introns: 27/2; 146/3; 268/1; 297/1; 349/2; 474/1; 530/2; 585/1; 680/2

Query Match 17.5%; Score 630; DB 2; Length 756;

Best Local Similarity 24.6%; Pred. No. 7.1e-37;

Matches 191; Conservative 132; Mismatches 291; Indels 162; Gaps 24;

```

Qy      30 FLAALSLPIAKTIGAIIMKSSIIHIERRFEISSLVGTFIDSGFEIGNLIVFVSYFGS 89
Db      7 FLIVFTVFLBESIGFMTSAVFIKQFLPSRLSGTMVSGADPAIYIPVIFTSYFGG 66
Qy      90 KLAHPRKLGIGCFIMGIGVLTALPHFPMGYRYSKETNIDS-----ENST 136
Db      67 KGNRRARWIGAGAILAIAPMLAASNFLLFVPEGANTTHIPSLAHQIDRIDVNGDVNST 126
Qy      137 STLSTCLINQIISLNRASPIV-----GKC----- 161
Db      127 EHL--WLRQIHPLIDPKQVVMNVYEGDEKSDLLQRYVEYCHYHKGSEICKRLKTYAEK 183

```

QY 162 -----CLKS-----GSTMIVFNGMLRGI 183  
DB 184 FPIITDAKSNVRAVVALPYGFCHSMNFVRAOHACKDDSTLGPFWMI--FGGLLVLAGV 241  
QY 184 GETPIVPLGLSYIDDPFAKEGHSILYLILNAMIIGIFTLGSJFSKMYVIGVYDLS 243  
DB 242 GRMPSPSLGLPLMDNDVKKNLPLYSFMFVKILGLVILVGGQANKLYDPN----P 237  
QY 244 TIRITPTDSRMVAGMMLNPLVSLGFSIISIPFFLPQTP---NKDQ-----KERRA 292  
DB 298 PGQITPLDPMWIGMNLGFLIFGLLPGSLVLYPFPSDDLDINABEPDODKVPKR- 356  
QY 293 SLSLHVLETDEKQDTANLNGKNITKAVTGFQSGKSLITNPLVYMFVLLTLQVSSY 352  
DB 357 --RLNLDRIHKDEQGNAL--MPETIMDKINDKATIANLVKRIFGAAGFRIIDVLAF 413  
QY 353 IGAFYVFKYVEOQYGGPSSKANILLGVITIPFASGMFLGYIILKKFKLNTVIAKFSQ 412  
DB 414 KGFVPOAKTLEHFGVPOYRIQRYIASGIVGFACGVILGISLMFELQG---RKAA 470  
QY 413 FTAVMSLSFYLLYF--FILCENKSVAG-LTMVTDGNNPVTSHRDVPLSYCNSDCNDES 468  
DB 471 WVAVCSILVALISFANGTVC--KSVIGQIGDQIKINGPV-----PDGCRDDCKCEQT 521  
QY 469 QMEPVCGNNGITITISPLAC-----KSSGKKEIVFPYNGSCLEVTGLQRRNYS 518  
DB 522 PLYPVCDVSGSAYSPCHACGAMKNSYIFDNTKATVDTLITENCQVD---NKVKE 576  
QY 519 AHLGECPRDACTRKFYFVAIQVILNLFSGALGTSHTMIVKIVQBELKSLALGFHSWV 578  
DB 577 VSRSYC-RTEECBERRFRFPMHQAFGAVFGGLGVVPCMLILVRAVPERHSVSLGNGFL 635  
QY 579 IRALGGLIADIVYQALIDTTCIKWSTNNCCTRGSCRTYNSPSRYLGL-LSMLAVSSL 637  
DB 636 VSLATLPSPIIWKIILWMSCLMK-KSCDSSGSCSYVD--TDRLRYRLHYICLAFISFL 693  
QY 638 VLTIIILY-----AMKKYOEKDINASENGWDEANLESINKKH 678  
DB 694 ISDVWVMYMARGLKLIBETPBDDKKEBETEKLSRKRESLKDAA--EKVERTAY 747

## RESULT 10

T27357

hypothetical protein Y70G10A.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T27357

Ri,lloyd, C.  
submitted to the EMBL Data Library, October 1998

A/Reference number: Z20354

A/Accession: T27357

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-690 &lt;MIL&gt;

A/Cross-references: UNIPROT:Q9XWCS; UNIPARC:UPI0000076276; EMBL:AL032660; PIDN:CAA21751.

A/Experimental source: clone Y70G10A

C/Genetics:

A/Genes: CRSP:Y70G10A.3

A/Map position: 3

A/Intons: 61/3; 84/2; 185/1; 250/2; 326/3; 375/1; 398/3; 439/2; 490/3; 628/1; 655/1

Query Match 17.2%; Score 618.5; DB 2; Length 690;

Best Local Similarity 26.8%; Pred. No. 4.2e-36;

Matches 169; Conservative 118; Mismatches 292; Indels 51; Gaps 17;

QY 28 KMFALALSL-SPIAKTIGALIMKSIHIERFEISSLVGFDGSRBIGMLVIVFVSX 86

DB 48 KCLLVVLGICAFIQSPVNAIPFVGSLTERFRKMTSTHGISSWYDFAVLVVPVCH 107

QY 87 FGSKLHRLPKLIGICPIMGIGVLTALPHFPMGYRYSKKTNISSSENSTSLTCLINO 146

DB 108 WGNNGHKGWIGVGMALGSLICALPHMMVDIY----HPDVNLDLTNQTD-FGQC----- 158

QY 147 ILSLNPASPEIVYGCKLCKSGSYMIVYFMGNMLRGIGETPIVPLGLSYIDDPFAKEGHS 206  
DB 159 ---ANRDBECAGKPHSHSNFNPFWMPI-LGOTLHGVSSTPLSISTTYMDENVSGKASP 214  
QY 207 LYLGIILNAMIIGIIGFTLGSJFSKMYVIGVYDLSIRIPTPTDSRMVAGMMLNPLVSG 266  
DB 215 VILAHAVLTSRQPVGVRAQGLNLVDPDRVE--RLPMERSDDRGMWVGITISS 272  
QY 267 LPSIISIPFF-FLPQTPNKKPOKERRASLSLHLETNDEKDTANLNGKNITKAVTGF 325  
DB 273 ISALMLAFILAFARLPEAKGRADVQNCVANDVAKAPRDLMLCPACVWK----- 327  
QY 326 PQSGKSLITNPLVYMFVLLTLQVSSYIAGFYVRYVQOQGGPESKANILLGVITIP 385  
DB 328 -----ILSNPFVLCIFGIFESIIINFPALFMPPILETLLSTNETLASYSVV--I 378  
QY 386 F--ASGMFLGYIILKKFKLNTVIAKFSCTAVMSLSFY--LLYFPILCENKSVAGLTMT 441  
DB 379 FAATVAVMGITIIIRQLKQVGMKMLIVCHVALLFTTGLSH---CPQKEFVGINIG 435  
QY 442 YDGNNEPVTSHRDVPLSYCNSDCNDESQMEPVCG-NGGITVISPCLACKSSGKKPIV 500  
DB 436 YEDLSIERKHDFSISSTCNADGHC-KMEWNPVCDRRTGHMYVSACHAGCTGRTTIDSSQ 494  
QY 501 FYNCSGLE-----YVGLQRRNYSAHLGECPRDACTRKFYFVAIQVILNLFSGALG 552  
DB 495 WSGCGCLTNSSTFNLHIGIKHPDVLNCGYQHD--CGRYEILMVTLPFIVVASFAAG 552  
QY 553 TSHVMLIVKIVQBELKSLALGFHSWVIRALGGILAPYEGALIDTTCIKWSTNNCCTRG 612  
DB 553 IPTQILMAVFPDQRTALGVNMFVRLLGFIPLGILIGIILIDFACLEWG-ESCGKAT 611  
QY 613 CRTNSTSPSRVYLGISMLRVSILVYII 642  
DB 612 CLVYDPEFKLSWITTGALVICKSLIATII 641

## RESULT 11

T21217

hypothetical protein F21G4.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T21217

R/Mortimore, B.  
submitted to the EMBL Data Library, October 1996

A/Reference number: Z19392

A/Accession: T21217

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-674 &lt;MIL&gt;

A/Cross-references: UNIPROT:Q93550; UNIPARC:UPI000007841C; EMBL:Z81016; PIDN:CAB02665.1;

A/Experimental source: clone F21G4

C/Genetics:

A/Genes: CRSP:F21G4.1

A/Map position: X

A/Intons: 58/3; 149/1; 181/3; 213/3; 286/1; 327/2; 457/2; 509/3; 608/1

Query Match 16.9%; Score 609; DB 2; Length 674;

Best Local Similarity 25.9%; Pred. No. 1.9e-35;

Matches 169; Conservative 117; Mismatches 283; Indels 84; Gaps 17;

QY 12 BAQPSSENKTRKCYCNGKMFALSLSPFINKTIGALIMKSIHIERFEISSLVGFDG 71

DB 63 ERTPLBGRKKYKRSINIFVLLMLVLAIVGQTYLGIV--GMLTTLERKRFBSSEKSGMLIS 120

QY 72 SFEIGNLIVIVSYFGSLARPKLIGICPIMGIGCVLTALPHFPMGYRYSKKTNISS 131

DB 121 LVDIHTHMAILLIGVISHYHLPRITGIGVILSSIMFALPLVLYFGTADYTOEQLMOK 180

QY 132 SENSTSTSLCLINQILSLNPASPEIVYGCKLCKSGSY--MMIVYFMGNMLRGIGETPIV 189

DB 181 KE-AVSVENSCDTN-----GRREISSQGBDCWRHEHHTHNAFTIILAFGLFPAGIFAPRN 234







```

QY 366 QYGPSSKANILGVITTRIPASGMFLGYIIKKK-----LNTWGIKAFSCPTAV 416
DB 358 ILGFSSGNLSLITGVGVVGVIVASIGIRISLGFENRKPSPMYFIIVGIAAGSSCL 417
QY 417 MSLSPFLVFLICENKSVAGLMTYDGNPNVTSHPDVP-----LSYCNDCNCDSEOWEP 472
DB 418 L-----IRCESLANYGV-----NVPSS---DLRYGALERTCSNCHD-SFPNP 457
QY 473 VCGNNG-IITYISPCLAGCKSSSGNK-KPIVYNGSCLVETGLONRNYSAHLCECPDPA 530
DB 458 VCSBDSKILFLSPCHAGCADMPKIGKAGSNMTWCG-----SKNTIVKGYC--DASC 508
QY 531 TRKFPFPAIVQVLNLFPSALGSGSHMLVKIVQPLSKLALG-----PHSNV 578
DB 509 QKQIYQFI-----IMFALS-----FCITPAVLAQSSSLRVNKHKRDHPTCGMLW 556
QY 579 IPALGGILAPIYFAGLIDPTTCIKSTNNGTSGSCRTYSTSPSRVYLGLSSMLRVSSLV 638
DB 557 MRLSGIRPAIVFGYIIVNCTWQ-KDCVSO-KQCFYASNLGNMFFPIAVKFTGCI 614
QY 639 IYIILYAMKKRYQEKD-INASENGSMDEANESL 673
DB 615 ILFLAAYC---YQESDKSNGKESCRTLETISEV 646

RESULT 14
B81254
probable transmembrane transport protein Cj1588c (imported) - Campylobacter jejuni (stra
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #ext_change 09-Jul-2004
C:Accession: B81254
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Kusterford, K.M.; Vanlille, A.; Whitehead, S.; Barre
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: B81254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-431 <PAR>
A:Cross-references: UNIPROT:Q9PM87; UNIPARC:UPI00000C1F8A; GB:AL139079; GB:AL111168; NID
A:Experimental source: serotype O2, strain NCTC 11168
A:Genetics:
A:Gene: Cj1588c
C:Superfamily: citrate utilization determinant

Query Match 4.9%; Score 178; DB 2; Length 431;
Best Local Similarity 22.9%; Pred. No. 4.4e-05;
Matches 88; Conservative 67; Mismatches 115; Indels 115; Gaps 18;

QY 71 GSFEIGNL---LVYFVSIFSGSKLRPKLIGICFIMIGCVLTALPHFPMGYRYSKET 127
DB 58 GAFAAGYLARPLGGIVMAHFGDKFGRKN---FWLSI--LLMVLPTVLAF----- 103
QY 128 NIDSENSSTLSTCLINQILSLNRASPEIVGKGLKSGSYMMIVFPMGNLRGISTP 187
DB 104 -----IPGETLGLAPVL-----ILRIFGIAIG----- 131
QY 188 IVPLGLSYIDPFAKGHSSLYGLIAMLAMIGPIIG---FTLGSLESKMYVDIGYDLST 244
DB 132 ELFGAMVFRBYCOEKOKAFPLSCINSAVALGILGSIYFLIINAFSIEIRIAY----- 186
QY 245 IRTPTDSRWGAMNLNPLVSGLSIISIPFPPLPOTPNKPQKERRASLSLHVLSTND 304
DB 187 -----AMRIAFVGGIFGIISYLRFRFQETVPFKOMKSSLSFPL----- 229
QY 305 KQOTANLTQGNKITKNTVGFQSFQFS--ILNPLVLMFVL-----TLQVSSYIGAF 357
DB 230 KD-----LFKEDIYVNL---FSSMMMTVLTGCVIVLMLPKMPSTLANSI--SG----- 275
QY 358 YVFKYVEQYQGPSSKANILGVITTRIPASGMFLGYIIKKPKLNTVGIASFCTAVM 417
DB 276 -----VEGSYLQ-----ILGILGIAL--GGAFM-GYLVDFK-----GLFKICIFPSLT 315

```

```

QY 418 SLSPFLVFLICENKSVAGLMTY 442
DB 316 FVFFSFLVYALYELKNLVLVCILY 340

RESULT 15
B84616
hypothetical protein Ab2g22730 (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #ext_change 09-Jul-2004
C:Accession: B84616
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.
enes, D.; Meriman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84616
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-507 <STO>
A:Cross-references: UNIPROT:Q9ZQ41; UNIPARC:UPI00000A2D4A; GB:AB002093; NID:G4314370; PI
A:Genetics:
A:Gene: Ab2g22730
A:Map position: 2

Query Match 4.9%; Score 178; DB 2; Length 507;
Best Local Similarity 19.7%; Pred. No. 5.3e-05;
Matches 98; Conservative 68; Mismatches 170; Indels 162; Gaps 19;

QY 6 HNKTRAEQPSNKKTRYNGL-KMFLALSLFIKTLGAILMKSIIHRRREISS 64
DB 63 YMDRAIASNVGNSITRSCNDGKCTLAT-----GIQGHFNLNF 102
QY 65 LVGFIDSGFEIGNLTVI-VFVYFSGSKLRPKLIGICFIMIGCVLTALPHFPMGYR 123
DB 103 EDGVLSSTFMGLLALSPFNL-----AKRLIGVLVWTI-AVL----- 142
QY 124 SKETNIDSENSSTLSTCLINQILSLNRASPEIVGKGLKSGSYMMIVFPMGNLRG 183
DB 143 -----GC-----GSSFAFWFIVLCMPFVG 162
QY 184 GRTPIVPLGLSYIDPFAKGHSSLYGLIAMLAMIGPIIGFTLGSLESKMYVDIGVDS 243
DB 163 GASRTSLAAPPIDNALPOEQKAMDLFRMCIPEGVALGIVGKGHF----- 213
QY 244 TIRITPTDSRWGAMNLNPLVSGLSIISIPFPPLPOTPNKPQKERRASLSLHVLSTND 303
DB 214 -----SMRYAFWGEAVLMAFPVVG---FLMKPLQLKGSRTLKNRNLQV-DNEI 259
QY 304 EKDQ-----TANLTQGNKITKNTVGFQSFQFSILTNPLVLMFVLTLQVSSYIGAF 358
DB 260 EHDQFEVSIETSKSSYANAVFSGFPAQDMVLYLKEKFFVNV-LGYVSYNFVIGAYS 318
QY 359 -----VFKYVEQYQGPSSKANILGVITTRIPASGMFLGYIIKK-----FYL 402
DB 319 WCPKAGNVIK-----MGRADMIRGAVTIIIGIYVTLGSGTILDRVATINAPFL 369
QY 403 --NTVGIKAFSCPTAVMSLSFY-----LLYFPLICENKSVAGLMTY 441
DB 370 LSGATFLGAVFCFTLFTLKSILGYFIALPALGELLVFATQAPVNAVYCLHCVKSLRPLSMA 429
QY 442 YDGNPNVTSHPDVP-----NVPSS---DLRYGALERTCSNCHD-SFPNP 457
DB 430 I---STVALHIFGDVPS 444

Search completed: March 14, 2006, 10:52:50
Job time : 49 secs

```